DOCUMENT RESUME

ED 449 220 TM 032 365

AUTHOR DeBray, Elizabeth; Parson, Gail; Woodworth, Katrina
TITLE Patterns of Response in Four High Schools under State Accountability Policies in Vermont and New York.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE 2000-04-00
NOTE 20p.; Paper presented at the Annual meeting of the American Educational Research Association (New Orleans, LA, April 24-28, 2000). This work is part of a larger, five-year study of the Consortium for Policy Research in Education on Accountability and the High Schools.
CONTRACT OERI-R308A60003
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Accountability; Disadvantaged Youth; *Educational Change; Educational Policy; *High Schools; *Low Achievement; Performance Factors; *Responses; State Legislation; *State Programs; Teacher Attitudes
IDENTIFIERS New York; Reform Efforts; Vermont

ABSTRACT

This paper examines the responses of four high schools to new accountability policies in two states, New York and Vermont. In each state, one school that was better positioned with respect to the policy and one that was a target of the policy were studied. New York's policy attaches high stakes for students though the Regents high school exit examination. In Vermont, school accountability is part of a broader policy of comprehensive finance equalization and a self-assessment action planning process. In both states theories of action assume that performance information from the accountability system will drive change. Schools were examined through interviews with key personnel, including teachers. The variations between the types of schools in response to the policies far exceeded the variation attributable to state policy design. In the high-performing schools, the policies in both cases put a spotlight on helping low-performing students. In the low-performing schools in both states there was a pattern of compliance without capacity with minimal meaningful or productive engagement on the part of the staff. The needs of the two low-performing schools suggest blind spots in the state policies with respect to theory, use of data, incentives, and short-term, school-level goals. The common patterns suggest that state policymakers need to rethink policy instruments for the lowest performing schools. (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

E. DeBray

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Patterns of Response in Four High Schools Under State Accountability Policies in Vermont and New York

Paper presented at the Annual Meeting of the American Educational Research Association,
April 2000
New Orleans, Louisiana

ED 449 220

Elizabeth DeBray*
Gail Parson
Katrina Woodworth

TM032365

This work is part of a larger, five-year study of the Consortium for Policy Research in Education on Accountability and the High Schools, supported by a grant (No. OERI-R308A60003) from the National Institute on Educational Governance, Finance, Policymaking and Management, Office of the Educational Research and Improvement, US Department of Education. The principal investigators are Richard Elmore and Leslie Santee Siskin, of Harvard University, and Martin Carnoy, of Stanford University. The views expressed in this paper are those of its authors and are not necessarily shared by the U.S. Department of Education, CPRE, or its institutional members.

* The three authors contributed equally to the writing of this paper.

BEST COPY AVAILABLE

This paper examines the responses of four high schools to new accountability policies in two states, New York and Vermont.[1] In each state we selected a school that was "better-positioned" with respect to the policy and a school that we considered to be a "target" of the policy.[2] The differences in these two state policies were dramatic; therefore, we anticipated that we would find important variation in school response between the schools in New York and the schools in Vermont. We found instead that the state policies interacted with existing school structures and norms to produce divergent responses by school *type*. The two better-positioned schools we looked at in each state went beyond compliance with policy requirements, launching coherent responses that exceeded the mandate, finding ways to use the policy to effectively enhance and drive their own missions. In contrast, the two target schools in each state also complied with the policies' technical requirements, but did not have the internal structures needed to translate the mandate into the kind of coherent action that could more effectively lead to the improvement of instruction and student achievement.

We found these contrasts to be compelling. This paper illustrates the differences in the stories of the four schools — one of each type in each state. Through an analysis of these schools, we explicate how and why the state policies landed so differently, with special focus on a) the degree of alignment between the states' accountability policies and the schools' internal accountability systems, and b) specific school-level activities in response to the policies.

There are several design differences between the two states' policies. While both centered on state testing systems, New York's policy attached high stakes for students — via the Regents high school exit exam. Until 1996, high school students in New York state earned either a "local diploma," after passing five Regents Competency Tests (RCTs), or a Regents diploma, which meant passing eight higher-level subject matter exams. In 1996, the Board of Regents announced that over the ensuing eight years, the competency tests would be phased out; all students would be required to pass five higher-level Regents exams to graduate (in English, math, global studies, U.S. history, and science). This meant that the graduating class of 2000 would be held accountable for passing the English examination and the class of 2001, for passing both English and math examinations. In contrast, there were no stakes attached to Vermont's New Standards

---

[1] The data collected for this paper was part of Phase 2 of fieldwork, in which a research team conducted interviews in twelve different high schools in four states, Texas, Vermont, Kentucky, and New York (the sample consists of three high schools per state). In this paper, we analyzed English and mathematics teacher interviews, including department chairs, and drew on notes from classroom observations conducted in those departments. We chose English and math because they were tested subjects, and could be compared across states. For each school, we also analyzed administrator interviews and interviews with selected teachers in special roles. This fieldwork was conducted over the course of several consecutive site visits in the school years 1998-99 and 1999-00.

[2] The design for the larger CPRE study called for selecting schools that were differently positioned with respect to the reform. The terms "target" and "better positioned" were applied to schools in reference to their position. A "target" school was not seen as successfully preparing its students (according to prior measures), but was also not identified as a candidate for state intervention; a "better positioned" school's prior performance was deemed adequate with regard to student performance on existing assessments. The rationale for selecting differently positioned schools in each of the four states was to explore the effect of uniform policies in schools with different histories. For a more detailed discussion of the project design, see Siskin and Lemons, "Internal and External Accountability and the Challenge of the High School," presented at the American Educational Research Association, April 2000.

Reference Exam (NSRE), except the printing of results in local newspapers. Neither state had clearly defined stakes for adults or a specified role for districts.

Vermont school accountability was part of a much broader policy, Act 60, passed by the Vermont legislature in 1998. Act 60 was a comprehensive finance equalization plan which included statewide testing, a process of school-level "action planning," and technical assistance for low-performing schools. Action planning was designed as a self-assessment *process* for capacity building at the organizational level, whereas New York policy focused only on measurable *outcomes*, i.e. individual students' passing scores. Action Plans were to include identified areas of improvement (in part, utilizing data from the NSRE) and specific goals. Schools were to submit plans to the district and state. Vermont offered no concrete performance targets that schools should aim for in their Action Plans.

Vermont's inclusion of action planning stood in contrast to New York's focus on "Regents tests for all," which was intended to be a lever for demanding high performance standards for all students and, as a result, for decreasing the statewide performance gap. Unlike in Vermont, New York state policymakers did not offer any specific guidance about *how* high schools should adjust to this significant policy shift. The details of implementation and organizational incentives were left entirely up to schools and districts. In both states, we studied the early implementation of these policies: the testing system was being phased in for juniors in New York, and the first cycle of testing and action planning was being carried out in Vermont.

In our examination of the schools' responses to the state policies, we focus first on the alignment between internal, school-level conceptions of accountability and the external, state-level accountability policies. In conducting these analyses, we relied on the working theory of school-site accountability developed by Abelmann, Elmore, and colleagues in the previous phase of this CPRE study of accountability. This theory posits that a school constructs a system of accountability out of the "relationships among three factors: individual conceptions of *responsibility*; shared *expectations* among school participants and stakeholders; and *internal and external accountability* mechanisms."[3] The theory assumes that there may be inconsistencies or complementarities in these relationships, and argues that internal accountability systems are likely to be stronger where there is alignment among these factors. Abelmann, Elmore, and colleagues developed this working theory based on the assumption that:

> The long-term fate of educational reform, as it is presently conceived, lies largely in th[e] tension between uniform requirements of external accountability systems and the particularities of real schools.[4]

With this in mind, in addition to expecting variation by state, we expected that school-level responses to the external accountability systems would depend on how well these policies

---

[3] Charles Abelmann, Richard Elmore, et al., *When Accountability Knocks, Will Anyone Answer?* (Philadelphia, Pa.: University of Pennsylvania, Consortium of Policy Research in Education, 1999), 3.

[4] Ibid, 1.

aligned with the schools' existing internal accountability systems. We found this to be the case in both states.

In the sections that follow, we examine these four schools, looking first at the "better-positioned" school, followed by the "target" school, in each state.

RIVERA HIGH SCHOOL, NEW YORK

Serving approximately 1,500 students, Rivera is a "screened" high school — that is, the school can be selective in choosing academically oriented freshmen. The student population is predominantly Latino (63 percent) and African-American (30 percent). Thirty percent of students receive free or reduced-price lunch. Historically, all students took Regents courses, and most sat for exams in math, science and English. Among 1999 graduates, 35 percent received Regents diplomas by passing all eight required exams. Therefore, Rivera was well-positioned to accept the state's challenge of getting all students to take and pass Regents exams because this has been its prior function.

Internal Conceptions of Accountability

Rivera educators were clear about their goal: to prepare students for college. As one English teacher said, "I'm not here to prepare them for their career. I'm here to prepare them for college... I want to teach them how to write analytically. And I think the Regents is a good assessment of it." The principal clearly stated that preparing students to pass the Regents had long been a schoolwide "reality":

> There's very...little flexibility in what we need to teach, if you care about your kids passing that test. ...[it's] the reality of this school, and now it's really a reality for everybody. I mean, it's been a reality for us for...many years, because we were always very...concerned about our data, and about the kids doing well here.

In this sense, Rivera is the school that best illustrates external accountability having been a factor in shaping the teachers' expectations in the years before the "Regents for all" policy took effect. Rivera's positioning with respect to the new policy reflects the fact that its prior function shaped an internal accountability system that is aligned with the new, external system.

The state testing mandate was readily embraced as the driving force behind the school's efforts, and was accepted as the measure against which Rivera educators judged their success. The principal spoke clearly about increasingly holding teachers accountable for student performance on the Regents. Although she did not have the authority to hire and fire, she used more informal means to hold teachers accountable for results:[5]

> We're all going to be looking at your results, and you're going to be held accountable...Can I fire teachers? No. Can I make them feel uncomfortable when they're not performing? Yes.

This translated into an internal accountability system that included publicizing individual teachers' Regents results. An English teacher explained:

---

[5] The feminine pronoun is used throughout this paper to ensure confidentiality.

Q. What are teachers actually held accountable for here?
A. Regents results certainly count. Those are public.
Q. And are they individually tied to teachers?
A. Oh, yeah.
Q. This group of students didn't do well and they all had Miss X.?
A. Absolutely. Everybody knows that. And not only in our own subject area, but there's like a school report card that comes out, and it's broken down by subject and then by individual class.

At Rivera, scores did not vanish into the organization, but were used.

Another math teacher talked about feeling personally responsible for these results in two ways:

> The principal and the [department chair] for mathematics look at the pass rates on the Regents exams as kind of a guideline of what's going on. And people whose pass rates are low are going to be scrutinized to see what they're doing, and what they can do differently. So I have an incentive to...get my students to do well on the exams...These kids come in expecting to learn things, and get ready for college. And I would feel remiss if I weren't supplying the best possible instruction for them.

What she believed she personally owed her students was supported by the school's internal accountability system. Teachers' feelings of personal responsibility formed a set of shared expectations that were reinforced internally by collaborative work and an active and focused administration. Teachers told us about analyzing test results within departments, as well as with the principal, who conferred with individual classroom teachers about their performance. All these elements played a strong role in establishing internal accountability:

> The administration in this school is very...very focused. They're very actively involved in the staff and the staff's performance...What I mean by focus, is that...we are all working towards the same goal...Everybody's pushing toward high achievement, in terms of state standardized exams. ...The first week that I taught in the school, the principal had already made an appointment, scheduled an appointment to sit in on one of my classes, just to be sure that I am following the standards of [Rivera], which is tremendous. I felt that that showed that even the principal was actively involved in every single one of his staff's performance, and how they're carrying through their objectives.

Clearly, the school had a strong internal accountability system that was aligned with the state's accountability system.

The Response to the Policy

We found a schoolwide consensus that Rivera was experiencing new pressures to improve its performance. Specifically, the responses involved using data to inform instruction, as described above, engaging in ongoing professional development, and adding instructional time before and after school, particularly for students most at risk of not passing the Regents.

Opportunities for professional development were varied. Most professional development was school-based, including formal faculty and department meetings, as well as study groups formed by teachers to discuss specific topics of interest. The principal described the array: mandated meetings twice a month, staff development days, subject area department conferences, and a forum she created, teacher study groups.

Department meetings covered a range of topics. The math department met in different configurations depending on what they were addressing. For example, members of the department met in sub-groups based on the courses they taught in order to share curriculum and assessment. Across the departments, the topics of rubrics and standards were integral to these conversations.

To achieve the aggregate goal of the "Regents for all" policy, teachers and administrators were paying increasing attention to those students most likely to fail. Math tutorial was available before school for those students most at risk of not passing the Sequential I exam (which every student eventually had to pass for graduation). The school offered courses in English, math and American history after school for Regents preparation, and there was a study skills course after school for ninth graders. The school had sought external funding from a foundation to support much of this after-school work, demonstrating their ability to increase capacity through external partners.

* * *

Rivera educators anticipated the challenge of getting all of its students to meet the state's standards by building on their existing capacity. New York's policy shift provided a galvanizing moment for staff to consider a variety of academic interventions for students. It is a goal that most believe they will meet. Robinson High School, our next case, was situated very differently from Rivera relative to the new state accountability system.

Robinson High School, New York

Robinson is a large, comprehensive urban high school. It serves 3,300 students in grades nine through twelve. Of those, 58 percent are African-American and 37 percent are Hispanic. Roughly 70 percent receive free lunch. The school is required by the district to enroll students with a broad range of academic skills (as measured by standardized tests administered in middle schools). While official dropout statistics at Robinson hover around five percent, there is a significant disparity between the number of ninth graders (over a thousand) compared to twelfth graders (under 400).

Internal Conceptions of Accountability

Robinson did not have a history of preparing students to earn Regents' diplomas: in recent years, roughly five percent of graduating seniors earned them. Instead, Robinson's prior function included the myriad goals of a comprehensive high school (i.e., general education, college prep, school-to-career) as well as preparing students for minimum competency exams (the RCT) which had been the state's bar for graduation until 1996. In this context, most characterized the new policy as a significant challenge in terms of both teachers' previous practices and students' skill levels.

Several teachers talked about the magnitude of this change by explaining how differently positioned this school was from those in other parts of the state. Many of the teachers' comments reflected pessimism about their students' readiness for Regents testing. One said:

"Our schools are not functioning and haven't been for a really long time. I'm sorry to see the students be penalized for that in the name of raising the standards." As another teacher indicated, educators often felt powerless to affect their students' chances: "I think it's ridiculous to expect kids who have been basically allowed to get to high school with inferior skills, to improve in their high school years." While these comments reflected teachers' senses of efficacy with respect to the policy, and their concerns about the stakes falling on students, many teachers agreed with the idea of raising standards. As one said: "Passing the Regents is something that's long overdue. It's not something that the kids cannot achieve... If we lift the standards and say, 'Let's aim for something else,' they will rise to the occasion."

Given the contrast between Robinson's prior function and the goal of the new policy, it is not surprising that we found little evidence of formal, internal accountability mechanisms that aligned with the new policy. In fact, teachers talked about the absence of any mechanisms (internal or external) aimed at holding them accountable for student performance:

> I'm not going to be accountable, nobody is. If they look at my class list and see that ninety-nine percent of my students fail, is anything going to happen to me? No.

One teacher acknowledged that if she were held accountable for the performance of her students, "I would lose my job overnight."

Because teachers' feelings of personal responsibility were not shaped by a clearly defined prior function or by formal accountability mechanisms, there was no coherent schoolwide system of accountability. Instead, educators at Robinson essentially operated in isolation. In the terms of Abelmann, Elmore, and colleagues' theory of school-site accountability, this school is a place where individual responsibility dominates and accountability is "atomized." In this kind of environment, "the school's conception of accountability collapses, by default, into individual teachers' conceptions of responsibility."[6]

Response to the Policy

The school and the departments followed a pattern of compliance with external requirements (i.e., the elimination of an RCT track). However, the school's response to the policy along three dimensions (teacher professional development, using student achievement data to set goals and monitor progress, and adding time and remedial instruction for students) was limited or non-existent. The absence of any school- or department-level response led to a range of individual-level responses and a district intervention.

The Lack of Schoolwide Response

Carrying on their tradition of operating in isolation, Robinson teachers did not participate in departmental professional development or engage in instructionally focused conversations that would help them prepare students for the exam. Among teachers who saw this lack of

---

[6] Charles Abelmann, Richard Elmore, et al., *When Accountability Knocks, Will Anyone Answer?* (Philadelphia, Pa.: University of Pennsylvania, Consortium of Policy Research in Education, 1999), 39-40.

professional development as problematic, at least one argued that support for this work should come from the state:

> What I really need from the state ... is some resources and training..: I feel ill-equipped... I feel as though I'm learning on the fly... It's not that we're unwilling to develop our own stuff, but a little more help would be welcome.

The absence of professional development aimed at helping teachers make instructional changes in alignment with the "Regents for All" policy extended to the use of data. In 1998-99, we did not hear about any use of data in preparing for the exams. Based on comments, such as the following made by the former math chair, it appeared that school-based administrators lacked the capacity or resources to make use of the data, and again they called for support from the state:

> Last year, [the state] did give us results... [But] it should have been done with support, training, supplemental material, extra resources, consultants, everything. If they thought this was really important, it should have been done with the teachers, not told to the teachers.

When we collected data in the fall of 1999, the results of the previous spring's Regents were in, but were not being used. In fact, English teachers gave us varying accounts of the overall percentage of students who passed. One teacher explained how teachers could learn about students' performance on the Regents: "I think if you asked and you inquired and you needed to know some information, yeah, sure, you can get the information." When we asked, many teachers said they did not know if the students in their classes had passed or not, or if the students they had taught the previous year had been successful.

In 1998-99, the most significant response to the policy came in the form of added time for instruction (i.e., before-school tutoring, additional courses). However, in the fall of 1999, teachers reported significant problems with "programming" — getting students into the right classes. These problems undermined efforts to organize students in preparation for the Regents. For example, many students were not placed in the right math classes by mid-semester. In English, a class that was established for seniors who had failed the Regents was filled with sophomores, juniors, and seniors (including some who may have already passed the English Regents).

Individual Responses and a District Intervention

Without a coherent school or departmental response to the policy, there was a range of teacher responses based upon their personal senses of responsibility and efficacy. We illustrate this range using English teachers' responses to questions about the implications of the "Regents for all" policy. One teacher, who indicated that she was not changing her teaching of tenth graders to align with the new Regents, explained why she threw away any state standards guides given to her:

> If we had much better students, then we might wish to focus on the details of, "Oh, we need to polish up this skill or that skill." But we're dealing with kids who are mostly drowning. When you're drowning, you don't discuss which life belt is better, you just throw out the nearest life belt.

In contrast, a member of the department who was convinced of the merits of the new test had attended several state-sponsored sessions on the test. Her response to the policy reflected her personal commitment to ensuring its success:

> I am truly convinced of the merits of this exam, and it's a very exciting time to change and to have the courage at the Empire State to say, "This is no longer a good exam," and to go forward and change.

Most teachers at Robinson fell somewhere in between these two.

While the district seemed to provide little support for a more coherent school-level response (i.e., in the form of information, professional development, or assistance with whole school change), there was a serious district intervention during our second year of data collection. In the aftermath of very low pass rates on the June, 1999, math Regents (estimated by teachers at four to five percent), the principal and math chair were both replaced. During the 1999-2000 school year, the new principal focused mostly on building staff morale and developing high-profile relationships with the community. The new math chair, who was very specific about her plans ("I'm very narrow-minded... My interest right now is to raise the results"), began to use data as a "reality check," straighten out programming problems, and initiate team-based work as a forum for course planning and sequencing.

* * *

Robinson's response to the new state accountability system illustrates what happens when an external policy lands in a school where a) student skill levels are significantly below the policy goal, and b) there is little organizational capacity, incentives, or support for coherent response to the policy. The lack of capacity to respond reflects the school's prior function, the lack of shared expectations among teachers with respect to their work, and the absence of any internal accountability mechanisms aligned with the external policy goals. It remains to be seen whether the district intervention will help to build school-level capacity for reform and establish a coherent internal system of accountability.

Glen Lake High School, Vermont

Glen Lake, the only high school in a district on the outskirts of a small city, serves approximately 840 students. The student population is homogeneous, almost entirely White. Approximately 13 percent of students are eligible for free or reduced-price lunch. Glen Lake is a school that has long emphasized attainment. Prior to the arrival of the New Standards Reference Exam (NSRE), Glen Lake was consistently among the top two high schools in the state on its standardized test scores. Seventy-five percent of 1999 graduates matriculated to college. Although it has historically been a high performer in the state, the new performance-based system has caused the school to look carefully at the performance of all of its students. The school's administration and academic departments have not merely accepted the challenge of improvement, but carried it forward in a concrete way.

Internal Conceptions of Accountability

Glen Lake staff talked about traditions in the building: community support for funding (the teachers are among the highest paid in the state), and a large number of veteran teachers who are both independent practitioners and experts in their subjects. In this environment, teachers held shared mutual expectations about being the best at what they do. Individual teachers felt responsible to parents to deliver a traditional, pre-college education, so there was high alignment between collective expectations and the staff's personal responsibility. These collective expectations were used to embrace external demands, such as the NSRE and action planning.

These collective expectations also helped to create a strong, albeit informal, internal accountability system. For example, at Glen Lake, participating in external initiatives was an expectation. According to the chair of the math department, who had been in that position for 33 years:

> We're involved in everything in this school. We just buy into everything that comes down the pike. We try to meet every state standard. We try to meet the needs of all the parents... We're just that kind of a school. We're not a school that very often says no. That kind of explains us.

In this context, we found surprisingly little resistance to curricular change. Being a teacher at Glen Lake seemed to be a stronger influence on collective expectations than teachers' individual practice, expertise or preferences. Since "being out in front" and "involved" in external demands is part of the school's culture, teachers were willing to make adjustments to further collective school priorities.

The school's internal expectations provided the incentive to align with the external requirements, more than any pressure from the state or district. Action planning was a school-level activity, subject so far to virtually no external stakes. The school board did not choose to review or approve the state-mandated Action Plan at all in the last cycle, and the state's only feedback was to ensure Glen Lake conformed precisely to a state rubric. And that, said the chair of the School Development Team (responsible for action planning), suited Glen Lake just fine:

> We're doing this because we think it's valuable for us. And we're going to keep doing it, even if the state backs off... We're conforming to the state series of boxes, but we have something we want to get out of it... I know there are other schools that are just filling the boxes just to get past this...

As exemplified here, the culture of the building, the tradition of success, and the pressures of Act 60 converged, and the teachers and administrators were eager to respond to the external changes when they easily might have coasted on their pre-existing successes without extreme pressure from the community to change.

The Response to the Policy

For teachers, who were accustomed to the school being perceived as a state academic leader, the spring 1998 NSRE test results were startling. Only 31 percent of tenth graders "achieved the standard" or "achieved the standard with honors" in the area of "reading for analysis and interpretation." In math, 26 percent of students taking the exam "achieved the standard" or achieved "with honors" in the area of problem solving.

In response, Glen Lake embraced the state-mandated action planning. Act 60 did not specify how action planning should be done, simply that every school must have an Action Plan on file with the district at the end of every school year. Glen Lake's decision to require every department to set data-driven goals was initially viewed as too ambitious by the state, but the principal maintained that this was how the school wanted to do it. In part, Glen Lake took the process seriously because of budget cuts. Under Act 60, the school district lost 2.2 million dollars through "recapture," or giving back to the state. In May 1999, the school board voted down a measure to replace $600,000 of that loss. A community that had offered its unwavering support for funding increases would no longer do so. With school board members under pressure to keep the budget down, the staff recognized that documenting performance was inevitable.

Before action planning was mandated, Glen Lake had an "Indicators of Success" committee, whose function was to begin to look at data and quantifiable goals for the departments. The principal was able to smoothly transform this group into the action planning team, comprised of various department chairs, the principal, and other volunteers. No department, including such areas as guidance and music, was exempt from action planning. We found that the school's academic departments embraced the process. The action planning team's job was to help departments define problems based on data, and set annual goals; these plans then became the school's Action Plan that was submitted to the district. The team leader compared the process to the "scientific method."

In both the English and mathematics departments we found evidence of strong alignment across three levels of goal-setting: the school-level Action Plans, the departmental plan, and individual teachers' goals for the classroom. We also found that the establishment of student achievement goals with respect to the state test led to changes in curriculum and pedagogy in both the mathematics and English departments. As described below, each department set goals, teachers changed practices, and students performed better on the test.

The mathematics department's goal was that the number of Glen Lake students who "achieved the standard" would increase by at least 20 percent on the next NSRE problem-solving section. The department easily met the goal. The 1999 results achieved a 60 percent increase in this area over the 1998 results. To achieve this result, mathematics teachers consistently told us that they were emphasizing "open-ended" questions more than they had in the past. Another adjustment to align practice with the exam was the approach to teaching algebra. The NSRE in math is given at the end of sophomore year, and tests algebra and some geometry concepts. The department, facing the challenge of how to help all students become adequately proficient in algebra, this year created a "Level Two" algebra sequence serving between a quarter and a third of the entering class. In Level Two, instruction was drawn out over two years, freshman and sophomore. In regular algebra, instruction was heavily focused on textbook exercises that involve graphing calculators and explanations of how answers were derived--signs of alignment with New Standards.

In Glen Lake's English department, full of independent practitioners who had worked there for twenty years or more, the curriculum had been notable for its emphasis on traditional literature, with a focus on novels. But the New Standards English exam results raised concern

that only about the top 25 percent of students were flourishing with this approach. The department, with the goal of exposing students to the variety of literary styles and genres that they would be asked to analyze on the state assessment, built a new ninth grade curriculum around a literary anthology with a broader selection. While the anthology did not completely replace the teaching of novels in freshman year, teachers paid greater attention to what was required by the exams.

In 1998-99, the English Department's goal was that 50 percent or more of tenth grade students would "achieve" or "achieve with honors" the reading skills on the New Standards exam. The chair hired a reading specialist as part of its Action Plan to assist students who were having the greatest difficulty, and she offered a variety of types of reading assistance across the curriculum. This investment paid off: the number of students reading for basic understanding increased by 15 percent and the number reading for analysis and interpretation increased by 29 percent.

* * *

Like Rivera, Glen Lake's compliance with the policy complemented their mission, or the school found ways to make it do so. They made changes in their curriculum to prepare students for the state exams, and made the action planning an exercise that was useful for the school. Thus, it was a case of a school's internal accountability mechanisms (i.e., using data to make instructional decisions) being reinforced by an external policy (action planning), and becoming even more highly aligned. At the same time, the assessment results caused the school to look carefully at which students were achieving, and make changes in teaching and curriculum in an effort to reach more of them. Much less internal-external alignment was evident in the other Vermont school in our study.

Garrison High School, Vermont

Garrison High School serves a small Vermont town of 6500, with significant populations of migrants and refugees (12 percent) and low-cost housing areas. The school has 220 students, sixty-four percent of whom receive free or reduced-price lunch. Per pupil expenditure at Garrison was $4500 in spring 1999, as compared to the Vermont average of $5600. While 97 percent of students graduate from Garrison, only 30 percent generally go to four-year colleges.

Internal Conceptions of Accountability

When Vermont's Act 60 and standards-based reform policies arrived on its doorstep, Garrison was struggling with serious communication, discipline and attendance problems. Administrators were preoccupied with crisis management and operations issues, resulting in a laissez faire approach to instruction, and a long history of only minimal demands placed on staff and students. Many members of the small, veteran staff had not been formally evaluated in years. They were too few in number to have functional departments, often a productive unit of accountability in larger schools.

Garrison's NSRE scores for 1997-98 were low. Only two percent of students met the standard in mathematical concepts and problem solving. Similarly, only nine percent of students

met the standard for reading analysis and interpretation. In response, Garrison sought and received technical assistance, and complied with Vermont's Act 60 and standards-based reform policies. In the following year, 1998-99, some scores rose. But with the exception of writing conventions, all the tested areas in math and English remained far below state averages. The state then raised the bar, and its lowest "requiring assistance" category expanded from four to thirty-nine schools. Garrison was one of them.

Weak Alignment of External and Internal Accountability

With its required testing, action planning, and standards-based curriculum, the state policy seemed to offer a springboard for schools to generate shared goals around which to align individual responsibility and internal accountability mechanisms. But as Abelmann, Elmore, and colleagues point out, external accountability systems like Vermont's policy "operate at the margins of powerful factors inside the school."[7]

"Powerful factors" were at work inside Garrison. It had an atomized culture, and no coherent system of internal accountability. Individual teachers' feelings of personal responsibility, idiosyncratic and widely varied, were not aligned in ways that created common expectations, nor were they shaped by informal or formal accountability mechanisms such as administrative evaluations.

And the leadership team itself was atomized. The principal and the curriculum coordinator each espoused belief in the importance of improved instructional practice focused on standards, but operated on separate tracks, with no consistent or unified expectations toward holding teachers accountable for student achievement. This schism exacerbated the lack of alignment between the external accountability policy and the school's inner workings. The curriculum coordinator had the requisite knowledge but no authority, while the principal's ability to recognize good practice was not bulwarked by a willingness to confront teachers in core areas who were not teaching to standards. By December of the second year, no teachers in the core tested areas, with some of the lowest scores in the state, had been evaluated.

The years of non-existent or weak evaluation mechanisms had taken their toll. Teachers felt little accountability to formal authority or to the community:

> *Q: Who do you feel accountable to?*
> A: I don't know…I don't feel anything from the outside…even when the test scores come out, there are kids that are getting this information, and so you're constantly trying to do better…And I think if the administrators didn't think I was doing okay, then I would feel like there was—maybe because they think I'm doing okay, that that's an accountability that just is sort of seamless and I don't think about it. Same thing with the parents.
> *Q: No news is good news?*
> A: Sort of…there are communities in this area where …the parents are just constantly in your face. "You got to do this, this, and that." That doesn't happen here, so you don't think of them as being your customer either. You think about getting these kids good jobs and good schools.

---

[7] Ibid, 38.

This sense of "no news is good news" reflected how weak the internal accountability system was at Garrison. With little consistent communication and evaluation, external demands for accountability could not take hold.

Individual Responsibility in a Culture of Low Expectations

The curriculum coordinator spoke often of teachers' low expectations for and beliefs about students. Teachers could not or would not make the connection to the need to change and improve their own practice, particularly in the absence of any real pressure to do so. Students were not expected to go to college here, said several teachers, "and, [parents] are not pushing their kids to do better either...that's the reality of it."

In a school with these generally low expectations, and an incoherent organizational response, any external initiative will devolve to individual responsibility. This became very clear as we spoke to teachers about their responses to the various initiatives the new policy had generated, specifically, the encouragement of curriculum alignment with state standards and standards-based teaching. As one math teacher explained, the curriculum coordinator led attempts to align standards with the curriculum, but there had been no assessment of what teachers were actually doing in their classrooms.

In the core area of English, the two veteran teachers (both with over twenty-five years in the school) acknowledged that they had been exposed to reform ideas about teaching and aligning curriculum to standards. They each said there was no real English curriculum, and they were aware of the school's low scores, but they weren't particularly disposed to fundamentally changing their practice, nor did they feel compelled or accountable for doing so. As one said:

> I really don't see a commitment to changing the typical courses in the typical high school merely to meet the standards... I don't sit down everyday and say, what standard am I teaching today? Which maybe I should be. But I don't.

Teachers felt that working with standards or changing their practice was a matter of individual discretion. While there were individual teachers with thoughtful and/or improving practice, this work was self-initiated and inconsistent across the school.

The Response to the Policy

At Garrison, professional development and the push for instructional improvement stopped at the classroom door, where it became an individual teacher's choice. Vermont's required action planning process had the same stunted quality. An official Action Plan was generated, but the Plan reflected the nature of the school: disjointed, with no internal accountability mechanisms that could actually begin to change the collective norms of the atomized culture.

The principal looked to the Action Plan as an opportunity to dovetail schoolwide initiatives, but the final product was a patchwork of visions and goals, with weak accountability mechanisms. The principal said the process began with a staff-wide "satisfaction survey," followed by work with a standing committee of representatives from the high school and middle school. The final product, she informed us, was written by an outside consultant. The teachers

did not have any ownership of this plan because, the curriculum coordinator said, "it was just done by a committee somewhere." Not surprisingly, teachers we interviewed referred very little to the Action Plan.

The final Action Plan focused on discipline, curriculum alignment with standards, and the improvement of math problem-solving skills, in response to low NSRE scores in that area. (New math texts were purchased with Act 60 money.) The abysmal scores in writing and reading were not addressed as an official focus, except as embedded in the goal of having a "coordinated, challenging" standards-based curriculum. There were no formal accountability measures in the plan. In short, the action planning process was a missed opportunity to create a cohesive response to problems in the school.

Professional development had little coherence as well. It was either personally motivated and individually sought out in isolation, or offered schoolwide, but it alternately featured different programs. In a school as atomized as Garrison, with little internal accountability, and few collective expectations, professional development — even the mandated implementation of state curriculum standards — did not "take hold" in any consistent way across the staff.

* * *

Garrison's story is one of a school whose disjointed leadership, atomized culture, and lack of coherent internal accountability mechanisms rendered it almost impervious to the intentions of Vermont policy. It did not respond to recognized needs or identified problems other than how it always had: in a compartmentalized, fragmented way, with leadership working at cross-purposes and understandings, all in a context of low student expectations.

In a state famous for its fierce commitment to local control around educational matters, the individual educator at Garrison was able to take local control to its logical extreme. Instructional improvement, aligning curriculum with standards, accountability for raising test scores — these were just options, choices — and student achievement was a sad testament to the choices many educators had made.

Conclusion

In New York and Vermont, both policies held the goal of equalizing performance over time, and addressing the problem of low-performing schools and students. Both states' theories of action[8] assume that "performance information from the accountability system will drive change in schools and districts. The theory of action is that the accountability system will provide feedback on school performance that will then be used in school improvement planning."[9] The key limitation of this theory, however, is existing local capacity for an alignment of internal and external accountability. As Massell explained, "performance data often are not transparent and

---

[8] Donald A. Schon, *The Reflective Practitioner: How Professionals Think in Action* (New York: Basic Books, 1983).

[9] Diane Massell, *State Strategies for Building Capacity in Education: Progress and Continuing Challenges* (Philadelphia, Pa.: University of Pennsylvania, Consortium for Policy Research in Education, 1998), 32.

readily understandable [and] educators often do not have the requisite knowledge and skills to translate them."[10] The mechanisms of this theory of action vary and look different once they are *in use* at the school level. For example, public reporting, which assumes that if local stakeholders are uncomfortable enough with low performance they will provide the necessary local-level pressure to bring about improvement, was the most salient feature of Vermont's accountability system at the school level. In New York, the new tests with high stakes for students were the most prominent feature of the state's accountability system from the school perspective, suggesting a simplified theory of action — that is, that policymakers posit that giving a single set of tests statewide, and placing stakes on students, will cause schools to improve.

Despite these differences in state accountability policies and mechanisms, we found that the variation between types of schools in response to the policies far exceeded the variation attributable to state policy design. However, before we describe this variation, we want to remind the reader that these findings are limited in at least two important ways. First, the fact that the accountability policies in Vermont or New York at the time of our study did not involve explicit rewards or sanctions for adults raises a question about whether our findings about school response would hold in states that do. For example, in another state in our study, Kentucky, the state's incentives are designed to foster school improvement relative to its prior performance levels, and the entire organization is rewarded or sanctioned based on this performance. Our findings about school type and response may differ greatly under such policy conditions, and we intend to extend this analysis about school type and state policy design in a future paper. Second, our findings about the similarities of response by school type are not generalizable to the broader population of high schools. In other words, we do not want to suggest that every low-performing school fails to use data well in planning, nor do we conclude that all better-positioned schools respond actively to accountability policies. Schools' responses to policies are dependent on a wide range of contingencies. Here, we note the similarities that were present in this small sample across similar types of schools operating under two very different sets of state incentives that shared the goal of spurring improvement in low-performing schools. We highlight these different patterns of response by school type:

- **High-performing schools:** The policy, whether New York's or Vermont's, reinforced the schools' identity, prior mission of attainment, and mechanisms of using data. The policies at work in these schools put a spotlight on helping low-performing students, which staff accepted as a worthwhile, significant challenge. Addressing the learning needs of these students with a coherent plan reinforced their sense of efficacy. Both Rivera and Glen Lake began to learn how to set short-term performance goals for making improvements in student achievement – especially by sharing performance data within the department and setting corresponding instructional goals. In both of these schools, the policy developed a momentum of its own that "fit" the schools' goals. As the head of action planning at Glen Lake put it, "We're doing this because we think it's valuable for us. And we're going to keep doing it, even if the state backs off."
- **Low-performing schools:** We identified a pattern of "compliance without capacity" in the lower-performing schools. Compliance was *pro forma*, with minimal meaningful or productive engagement on the part of the staff. In these schools, other "survival" needs often eclipsed new,

[10] Ibid, 33.

external demands. At Robinson, for instance, teachers reported that poor course scheduling and sequencing for students was a barrier to tackling any further challenges. At Garrison, serious discipline and attendance problems were a daily reality. While both schools managed to comply with the policy, they failed to use it as a vehicle for instructional improvement. This failure to respond coherently was due, in large part, to the absence of collaborative structures and routines. Both Garrison and Robinson were responding to policies that did not specify short-term performance targets or achievement goals. In New York, Robinson was faced with the challenge of having to climb a very steep curve of getting all students to meet the standard, which overwhelmed them. At Garrison, the state had not specified any incremental achievement targets, so they selected "areas for improvement," but were not under any pressure to quantify by how much they would improve.

The needs of the low-performing schools in our small sample suggest several "blind spots" of these two state policies, as we see them. These can be thought of as ways that the state policies were under-specified with regard to improving student performance in all schools.

1. **The theory**. Policies are landing in high schools that have their own, particular internal accountability structures. If the elements of internal structure and external accountability systems are not aligned, then the policies may not pierce the instructional core and foster improvement. For instance, at Garrison, neither the presence of a common assessment nor the requirement of a planning process was able to overcome the inertia of the atomized internal structure of the school.

2. **Use of data**. Because of this lack of internal structure, schools may not have internal mechanisms for how to use the data in planning, or even understand its consequences for the school. As Massell noted above, it is an assumption that simply because a school receives results from a state assessment, it will know how to respond to it, and will possess the knowledge, skills and resources to respond. This also is a lot to expect from schools concerned with "survival," as we observed earlier.

3. **Incentives**. Using data may in turn generate an incentive problem: results may reinforce educators' beliefs about student abilities to learn, or their own efficacy to turn the problem around, which can perpetuate a cycle of even lower expectations. Examples of this scenario in our schools are abundant. In New York, the theory assumes that giving the Regents with stakes for students in and of itself will be an incentive for improvement. We wonder, when performance scores are as abysmal as they were in basic Regents math at Robinson last year, whether this may move from serving as an incentive, to instead, serving as reinforcement of blame and helplessness. At that school, a new math chair is arriving in the wake of a 95 percent failure rate on the basic math Regents last spring. The policy offers her nothing in terms of support, incentives, or models, for addressing instructional change. She arrived in a school absent collaboration or a functional internal accountability structure, and faced a policy that neither acknowledges nor supported her in dealing with that.

4. **Short-term, school-level goals for continuous improvement**. While the state policies envision improved long-term outcomes for schools, neither addressed how a low-performing school would set *short-term* performance goals. If Vermont's goal, for example, is for low-performing schools to tackle continuous improvement, it may have to mandate some reasonable performance target, or monitor schools in their selection of reasonable goals. Should Garrison, a school with only 33 percent scoring "acceptable" in reading interpretation and analysis, be permitted to submit an Action Plan that does not address literacy?

In addition to these policy shortcomings, we have two closing observations about each state policy in particular.

In New York, the policy assumes that students will be motivated to re-take failed exams repeatedly, and that teachers will be motivated to engage in endless remediation. Teachers are the

reform's implicitly relied-upon partners. Their motivation and effort are assumed in the state's additive equation about how it will work to keep students in high school for as long as it takes them to pass. The teachers are simply present, and presumably (since no state policies have provided otherwise) will be willing to commit extra time to teach the same content repeatedly in any way they know. Rhetoric at the state level has emphasized that students will have numerous chances to keep trying to attain a diploma if they fail; if they will not give up, they will have more than one chance to re-take an exam. Robinson's situation calls these assumptions into question. Will teachers there be "called" to repeatedly deliver remedial instruction to students who did not meet the state standard the year before? What are their incentives?

In Vermont, the policy assumes that giving all schools latitude to develop and execute an Action Plan that "makes sense" for their community will be fruitful, that local control will become part of the planning process that drives improvement. Act 60 assumes that test scores are an incentive; and that requiring action planning will result in productive, collaborative identification of priorities and goal-setting. Garrison's experience demonstrates how problematic these assumptions are. The administration drew up an Action Plan that failed to address a major area of weakness: literacy. The planning process involved minimal buy-in of teachers, with no apparent community activism nor stakeholders scrutinizing achievement results. In a deeply atomized school, the teachers felt they had a choice of whether or not to participate in any kind of improvement, let alone change their practice. The state was able to document schools' achievement deficits, but did not help to set concrete, realistic goals to strive for in the short run.

The common patterns of response we identified in these high schools indicate that state policymakers need to re-think policy instruments for the lowest-performing schools. Further, high schools' internal accountability structures are complex terrain, and how schools navigate external policy requirements is quite dependent on it. The response of the schools in this small sample points to the potential for the policies to increase the achievement gap within each state, exacerbating differences in school performance.

Works Cited:

Abelmann, Charles, Elmore, Richard et al. 1999. *When Accountability Knocks, Will Anyone Answer?* Philadelphia, PA: University of Pennsylvania, Consortium for Policy Research in Education.

Massell, Diane. 1998. *State Strategies for Building Capacity in Education: Progress and Continuing Challenges.* Philadelphia, PA: University of Pennsylvania, Consortium for Policy Research in Education.

Schön, Donald A. 1983. *The Reflective Practitioner: How Professionals Think in Action.* New York: Basic Books.

Siskin, Leslie S. and Lemons, Richard. 2000. *Internal and External Accountability and the Challenge of the High School.* Paper presented at the annual meeting of the American Educational Research Association, April 2000, New Orleans, Louisiana.